Asymptotic Rate of Decrease of Genetic Variability in Subdivided Populations (生物モデルの数学)

AUTHOR(S):

MARUYAMA, TAKEO

CITATION:

MARUYAMA, TAKEO. Asymptotic Rate of Decrease of Genetic Variability in Subdivided Populations (生物モデルの数学). 数理解析研究所講究録 1973, 174: 247-276

ISSUE DATE:

1973-03

URL:

http://hdl.handle.net/2433/107056

数理解析研究所講究録
174巻 1973年 247-276

A stochastic problem in population genetics: Asymptotic rate of decrease of genetic variability in subdivided populations*

TAKEO MARUYAMA**
National Institute of Genetics, Mishima, Japan

1. Introduction

The rate at which a population approaches homozygosity is an important number in population genetics. In particular, the asymptotic rate is related to the amount of genetic variability maintained in a population and to the speed of gene substitution. This problem was first investigated by Fisher (1922, 1930) and by Wright (1931). Wright found the general result that the rate in a randomly mating population of N diploid individuals is equal to 1/2N. Therefore, if a population is polymorphic at a locus initially, the probability of still being polymorphic at the t-th generation is proportional to $(1 - \frac{1}{2N})^t$, assuming no mutation or immigration from outside of the population. When a population is divided into partially isolated colonies or lines, or when there is

---

* Contribution No. 816 from the National Institute of Genetics, Mishima, Shizuoka-ken, 411 Japan. Aided in part by a Grant-in-Aid from the Ministry of Education, Japan.

**国立遺伝学研究所　丸山毅夫

assortative mating, the problem becomes more difficult. Robertson (1964) clarified the effect of non-random mating on this rate in a general theorem. In previous papers, I have studied the rate for a population linearly subdivided into colonies, Maruyama (1970a, 1970b). Letting n be the number of colonies, m be the migration rate between adjacent colonies, and N be the number of individuals in a single colony, the rate is equal to the smaller of $10m/2n^2$ and $1/2Nn$. The latter value would be the rate if the whole population were combined into a single population of size nN. The former value, $10m/2n^2$, is independent of the total number of individuals in the population.

Populations occupying two-dimensional habitats are biologically more interesting cases. In this paper I want to investigate such cases. First we study cases of torus-like lattice structure and then cases with rectangular habitats.

2. Torus-like space

Consider a torus-like lattice of n x n points, i. e., the direct product of two circular lattices with n points. Assume that a population consists of n^2 colonies and that each colony occupies a grid point of the lattice. Assume furthermore that there are N diploid individuals in each colony.

Let m be the rate at which a colony receives immigrants from its four neighboring colonies in one generation. It receives

immigrants from each at a rate of m/4 per generation. Let $f_{ij}^{(t)}$ be the probability that two randomly chosen homologous genes, one each from colonies which are i-steps apart in one dimension and j-steps apart in the other, are identical by descent. Let $f_{00}^{(t)}$ be the probability of identity by descent when the two genes are sampled from a single colony without replacement. The recurrence relations for the $f_{ij}^{(t)}$ are

$$(2\text{-}1) \qquad f_{ij}^{(t+\Delta)} = \left\{(1 - m)^2 + \frac{m^2}{4}\right\}\phi_{ij} + (1 - m)\frac{m}{2}\left\{\phi_{i-1,j} + \phi_{i+1,j} + \phi_{i,j-1} + \phi_{i,j+1}\right\} + \frac{m^2}{8}\left\{\phi_{i-1,j-1} + \phi_{i-1,j+1} + \phi_{i+1,j-1} + \phi_{i+1,j+1}\right\} + \frac{m^2}{16}\left\{\phi_{i-2,j} + \phi_{i+2,j} + \phi_{i,j-2} + \phi_{i,j+2}\right\},$$

where $\phi_{00} = (1 - \frac{1}{2N})f_{00}^{(t)} + \frac{1}{2N}$ and $\phi_{ij} = f_{ij}^{(t)}$ for i and $j \neq 0$, and where Δ is the length in time of one generation. Now let

$$(2\text{-}2) \qquad h_{ij}^{(t)} = 1 - f_{ij}^{(t)}.$$

Assuming that reintroduction of lost alleles into the population by mutation or migration does not occur or occurs at a negligibly small rate, we are interested in determining the asymptotic rate

at which the $h_{ij}^{(t)}$ approach zero. If we let $H^{(t)}$ be the matrix $[h_{ij}^{(t)}]_{nxn}$, we want to determine a constant λ such that

$$(2\text{-}3) \qquad H^{(t+\Delta)} = \lambda H^{(t)}$$

or

$$(2\text{-}3') \qquad H^{(t+\Delta)} - H^{(t)} = (\lambda - 1)H^{(t)}$$

for large t. In order to determine λ, we will derive a partial differential equation from (2-1). First let us define a function

$$(2\text{-}4) \qquad h(t, \tfrac{i}{n}, \tfrac{j}{n}) = h_{ij}^{(t)}.$$

Although the above h(t, x, y) is defined for $0 < x, y < 1$, this function can be extended onto the whole x-y plane as a doubly periodic function of period 1.

From the recurrence relation (2-1) and definitions (2-2) and (2-4), we have

$$\begin{aligned}(2\text{-}5) \qquad h(t+\Delta, x, y) - h(t, x, y) = \tfrac{1}{2}(1 - m)^2 \Big\{ & h(t, x+\tfrac{1}{n}, y) \\ & + h(t, x-\tfrac{1}{n}, y) - 2h(t, x, y) + h(t, x, y+\tfrac{1}{n}) \\ & + h(t, x, y-\tfrac{1}{n}) - 2h(t, x, y) \Big\} + \tfrac{m^2}{16} \Big\{ h(t, x+\tfrac{2}{n}, y)\end{aligned}$$

$$+ h(t, x-\frac{2}{n}, y) - 2h(t, x, y) + h(t, x, y+\frac{2}{n})$$

$$+ h(t, x, y-\frac{2}{n}) - 2h(t, x, y) + h(t, x-\frac{1}{n}, y+\frac{1}{n})$$

$$+ h(t, x+\frac{1}{n}, y+\frac{1}{n}) - 2h(t, x, y) + h(t, x+\frac{1}{n}, y-\frac{1}{n})$$

$$+ 2h(t, x-\frac{1}{n}, y-\frac{1}{n}) - 2h(t, x, y)\Big\} + h(t, x-\frac{1}{n}, y+\frac{1}{n})$$

$$+ h(t, x-\frac{1}{n}, y-\frac{1}{n}) - 2h(t, x, y) + h(t, x+\frac{1}{n}, y+\frac{1}{n})$$

$$+ h(t, x+\frac{1}{n}, y-\frac{1}{n}) - 2h(t, x, y)$$

If we let n get large and Δ small, keeping

$$(2\text{-}6) \qquad \frac{\Delta}{\left(\frac{1}{n}\right)^2} = 1$$

We have the differential equation

$$(2\text{-}7) \qquad \frac{\partial h}{\partial t} = \frac{m}{2}\left(1 - \frac{m}{4}\right)\left\{\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right\} - \frac{\delta(x)\delta(y)}{2N} h$$

where $\delta(\cdot)$ is Dirac's delta function. The left side of (2-7) is equivalent to the left side of (2-3'). For fixed and small values of N and m, if we let n get indefinitely large, $h(t, 0, 0)/h(t, \frac{1}{2}, \frac{1}{2}) \approx 0$, because local homozygosity proceeds much faster than that

in the entire population. Under such circumstance, the boundary condition to be imposed on the first eigenfunction of the right side of (2-7) is

$$(2\text{-}8) \qquad \phi(0, 0) = \phi(1, 0) = \phi(0, 1) = \phi(1, 1) = 0$$

and $\phi(x, y)$ must be a doubly periodic function. Therefore the operator for which we seek the dominant eigenvalue is

$$L = \left\{ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right\}.$$

Since this operator with the above boundary condition has a unique system of eigenfunctions, if we let ϕ be the first eigenfunction, we have

$$L\phi = \left\{ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right\}\phi = \alpha\phi$$

where α is a constant (cf. Maruyama, 1970a). Therefore, with small N and m, when n gets sufficiently large, there is a constant α such that $\alpha\frac{m}{2}(1 - \frac{m}{4})$ is the dominant eigenvalue of $\frac{m}{2}(1 - \frac{m}{4})L$. This together with (2-6) implies that the dominant eigenvalue we seek is

$$(2\text{-}9) \qquad \lambda = 1 - \frac{1}{n^2}\frac{m}{2}(1 - \frac{m}{4})\alpha.$$

Since α in the right side of (2-9) is independent of n and m, it can be determined by a simple numerical calculation. Though this is true for all the eigenvalues, it is particularly easy to evaluate the dominant eigenvalue.

Let R be the circular matrix of order n, i. e.,

(2-10) $$\begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{bmatrix}_{nxn}$$, where all unspecified elements are zero,

and G be the matrix

(2-11) $$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{nxn}$$, where all elements are zero except the first.

Now let

(2-12) $$M = (1 - \frac{m}{2})I + \frac{m}{4}(R + R^{-1}).$$

Then the recurrence relation for $H^{(t)}$ can be expressed as

(2-13) $$H^{(t+\Delta)} = M[H^{(t)} - \frac{h_{00}^{(t)}}{2N}]M = MH^{(t)}M - \frac{h_{00}^{(t)}}{2N} MGM.$$

For sufficiently large n,

$$(2\text{-}14) \qquad H^{(t+\Delta)} = M[H^{(t)} - \frac{h_{00}^{(t)}}{2N} G]M = \left\{1 - \alpha\frac{m}{2}(1 - \frac{m}{4})\right\} H^{(t)}$$

where α is the constant to be determined. To evaluate α, we use the so-called power method. We start the iteration (2-13) with an arbitrary initial matrix, e. g., a matrix with each element equal to 1, and we normalize the matrix before each iteration. After a sufficient number of iterations, the matrix remains unchanged by the iteration, except for being multiplied by the constant λ. A few cases of numerical calculations are given in Table 1.

Table 1

As shown in Table 1, the α in (2-14) is nearly independent of n or m. Therefore we have

$$(2\text{-}15) \qquad \lambda \approx 1 - \frac{1}{n^2}\frac{m}{2}(1 - \frac{m}{4})$$

if n is large and N and m are small.

In the above analysis, we have assumed that $\phi(0, 0) = \cdots = \phi(1, 1) \approx 0$ by assuming that m and N are small. Let us now consider the other extreme situation, in which $N = \infty$. Then the right side of (2-7) becomes

$$(2\text{-}16) \qquad \frac{m}{2}(1 - \frac{m}{4})\left\{\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right\}$$

and the dominant eigenvalue is unity. A complete spectral analysis for (2-16) is possible. Now consider the right side of (2-7) with a finite value of N as a perturbed form of the ideal situation (2-16), perturbed by $-\delta(x)\delta(y)/2N$. Then we can apply the perturbation method to obtain an approximation for the eigenvalue of (2-7). The second order approximation turns out to be

$$(2\text{-}17) \qquad \lambda = 1 - \frac{1}{2Nn^2} + \frac{1}{(2N)^2 1.2n^2 m} .$$

See Kato (1966), for the perturbation theory. Thus, if N is large, as a first order approximation the largest eigenvalue of (2-7) is the same as that for a randomly mating population of size Nn^2. Table 2 shows some examples with large mN.

Table 2

For large n and small m, formula (2-15) is valid, and formula (2-17) is valid for large N. Transition from one formula to the other occurs at

$$(2\text{-}18) \qquad \frac{1}{2Nn^2} = \frac{m}{2n^2} \quad \text{or} \quad 1 = mN.$$

This is shown in Figure 1. In the figure, the ordinate represents the value of $(1 - \lambda)/(m/2n^2)$ and the abscissa represents the

256

Value of mN. The lines represent the values from the approximation formulas and the dots represent the exact values obtained numerically.

Figure 1

From this numerical evidence together with the analysis developed in this section, we may conclude that approximation formula (2-15) is valid for situations where $mN < 1$ holds, while approximation formula (2-17) is valid for situations where the inequality is reversed. A set of parameters for which $mN = 1$ holds represents a transition point in that the validity of one approximation formula breaks down at this point and the other formula becomes more valid from this point on. The numerical examples indicate that the transition behavior of the eigenvalues, in the sense discussed above, takes place rather rapidly.

It is important to note that in the two-dimensional cases whether the eigenvalue is given by (2-15) or (2-17) depends only on the value of mN (the number of migrating individuals), but nearly independent of the habitat size, while this matter depends on both mN and the habitat size in the one-dimensional cases, (see Maruyama 1970a, 1971a).

3. Square plane habitat

In this section, we consider a population subdivided into colonies which occupy a square lattice of n by n colonies. Thus, unlike the cases in the previous section, the habitat dealt with in this section has boundaries. We assume also that the total migration rate into a single colony from its neighboring four colonies is m, the migration rate from each colony being m/4. A colony located at the boundary receives immigrants from the colony

on the side away from the boundary at the rate m/2, which is twice the rate for a colony not at a boundary.

Let us designate the position of a colony by a pair of integers (i, j), where $1 \leq i, j \leq n$. Let $f^{(t)}_{i_1 j_1 i_2 j_2}$ be the probability of two homologous genes taken from colonies (i_1, j_1) and (i_2, j_2) being identical by descent, and let

$$h^{(t)}_{i_1 j_1 i_2 j_2} = 1 - f^{(t)}_{i_1 j_1 i_2 j_2}.$$

The $h^{(t)}_{i_1 j_1 i_2 j_2}$ satisfy a recurrence equation similar to (2-5), and, as the limit when $n \to \infty$, we have

$$\frac{\partial h}{\partial t} = \frac{m}{2}\left(1 - \frac{m}{4}\right) \sum_{i=1}^{4} \frac{\partial^2 h}{\partial x_i^2}.$$

If we apply the transformation

$$z = x_1 - x_2$$
$$w = y_1 - y_2$$

and ignore the difference in h(t, x, y, x, y) for large t, we have

$$\frac{\partial h}{\partial t} = m\left(1 - \frac{m}{4}\right)\left\{\frac{\partial^2 h}{\partial z^2} + \frac{\partial^2 h}{\partial w^2}\right\},$$

the same equation as (2-7), except for a factor of 2. With large

n and small mN, the boundary condition equivalent to (2-8) is now

$$\phi(0, 0) = \phi(0, 2) = \phi(2, 0) = \phi(2, 2) = 0.$$

Instead of $\Delta n^2 = 1$ of (2-6), we now let $\Delta(2n)^2 = 1$. We then have

$$(3\text{-}1) \qquad \lambda = 1 - \frac{1}{4n^2} m(1 - \frac{m}{4})\alpha \approx 1 - \frac{1}{4n^2} m(1 - \frac{m}{4})$$

so that $1 - \lambda$ is half of the $1 - \lambda$ given in (2-15).

For large mN, by the perturbation method, we obtain

$$(3\text{-}2) \qquad \lambda \approx 1 - \frac{1}{2Nn^2} + \frac{1}{(2N)^2 2.4n^2 m}$$

The transition from (3-1) to (3-2) occurs at $mN = 2$.

4. Non-square habitats and habitats of general migration pattern.

In the preceding sections, the dominant eigenvalues for square habitats were given in simple forms, although these were approximations, (2-15), (2-17), (3-1) and (3-2). The problem for a non-square rectangular habitat seems to be more complicated. In this section, we will consider the problem for such habitats and for populations with a more general migration pattern. We take

an entirely different approach to these cases.

Consider a population consisting of $n_1 \times n_2$ colonies arranged on a n_1 by n_2 torus-like lattice. We now allow mutation to occur and assume that every mutant is new to the population. Let u be the mutation rate per gene per generation. Notice that, although mutations occur, if u is sufficiently small, there will be at most a few different alleles at a locus, and the whole population will be homallelic most of the time. Therefore if we can get a complete analysis for such a steady state, we will know the transient behavior of a particular mutant gene. Let f_{ij} be the same probability as f_{ij} in section 2, except that it is for the steady state and therefore is without the superscript (t). Then the matrix $F = [f_{ij}]_{n_1 x n_2}$ satisfies

$$(4\text{-}1) \qquad F - (1-u)^2 M_1 F M_2 = (1-u)^2 (1-f_{00})(2N)^{-1} M_1 G M_2$$

where M_i is the matrix (2-12) in which R is the circular matrix (2-10) of order n_i, and G is the matrix (2-11). The explicit solution (f_{ij}) of (4-1) is

$$(4\text{-}2) \qquad f_{ij} = \frac{1-f_{00}}{2Nn_1n_2} \sum_{k=0}^{\left[\frac{n_1}{2}\right]} \sum_{\ell=0}^{\left[\frac{n_2}{2}\right]} \frac{\Delta_k^{(1)} \Delta_\ell^{(2)} \xi_{k\ell} \cos \frac{2\pi ik}{n_1} \cos \frac{2\pi j\ell}{n_2}}{1-(1-u)^2 \xi_{k\ell}}$$

260

where $\xi_{k\ell} = \left[1 - \frac{m}{2}(1 - \cos\frac{2\pi k}{n_1})\right]^2 \left[1 - \frac{m}{2}(1 - \cos\frac{2\pi \ell}{n_2})\right]^2$,

$\Delta_0^{(i)} = \Delta_{\frac{n_i}{2}}^{(i)} = 1$ and $\Delta_k^{(i)} = 2$ otherwise, and [x] is the greatest integer $\leq x$. From (4-2), f_{00} can be determined explicitly;

$$(4\text{-}3) \qquad f_{00} = \frac{S}{2Nn_1n_2 + S}$$

where

$$S = \sum_{k=0}^{\left[\frac{n_1}{2}\right]} \sum_{\ell=0}^{\left[\frac{n_2}{2}\right]} \frac{\Delta_k^{(1)}\Delta_k^{(2)}\xi_{k\ell}}{1 - (1 - u)^2\xi_{k\ell}} .$$

For the derivation of (4-1) to (4-3), see Maruyama (1971). Note that

$$(4\text{-}4) \qquad \bar{f} = \frac{1}{n_1n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} f_{ij} = \frac{1 - f_{00}}{2Nn_1n_2[1 - (1 - u)^2]} .$$

Robertson (1964) has shown that $1 - \bar{f}$ decreases, as a result of random genetic drift, at the rate

$$(4\text{-}5) \qquad 1 - \lambda = \frac{1}{2N_T} \frac{1 - f_{00}}{1 - \bar{f}}$$

where $N_T = Nn_1n_2$ = the total population size. In particular, $f_{00} = \bar{f}$ in a random mating population and therefore $1 - \lambda = 1/2N_T$,

the well-known result of Wright (1931). Robertson's $\bar{f}$ and f_{00}, and the $\bar{f}$ and f_{00} of (4-3) and (4-4) are different, but if the mutation rate u is sufficiently small, they are approximately the same. Therefore we can use formula (4-5), together with formulas (4-3) and (4-4) to calculate the rate of decrease. Several examples of the value of $1 - \lambda$ calculated by (4-5) and the exact value obtained by the iteration method are compared in Table 3.

Table 3

So far we have restricted the migration to that occurring between geographically adjacent colonies. Let us now relax this restriction and allow a more general migration pattern. Let m_{xy} be the migration rate between colonies separated x steps and y steps along the first and the second coordinate axis respectively, i. e., the colony at (i, j) receives migration from the colony at (i+x, j+y) at the rate $m_{xy}/4$, and at the same rate from the colonies at (i+x, j-y), at (i-x, j+y) and at (i-x, j-y). In particular, $m_{10} = m_{01}$ corresponds to the m/2 of the preceding analysis. With this general migration pattern, the only modification needed in formula (4-2) and (4-3) is

$$\xi_{k\ell} = \prod_{x,y} \left[1 - \frac{m_{xy}}{2} \left(1 - \cos \frac{2\pi xk}{n_1} \cos \frac{2\pi y\ell}{n_2} \right) \right]^2$$

where $\prod_{x,y}$ indicates the product involving all possible combinations of x and y.

The calculation of $1 - \lambda$ from (4-5) involves the calculation of f_{00} and $\bar{f}$ which require a computer. Therefore it is not as simple as those in the previous sections. However, in this way, we can calculate the value of $1 - \lambda$ for a habitat whose shape is nearly arbitrary, and for cases with parameters having intermediate values. Formulas for f_{00} and $\bar{f}$ in rectangular plane habitats are given in Maruyama (1971).

5. Eigenfunction

The eigenfunctions, particulary that associated with the dominant eigenvalue, are of great importance. It seems difficult to obtain the exact eigenfunctions. However, we have the $[f_{ij}]$ in the steady state and we may use these to get an approximation to the eigenfunction. With sufficiently small u, the eigenfunction H associated with the dominant eigenvalue is given by

$$(5\text{-}1) \qquad H = [h_{ij}] \approx c[1 - f_{ij}]$$

where the f_{ij} are given by (4-3), and where c is a normalizing constant. In Table 4, the eigenfunction H given by (5-1) is compared with the exact eigenfunction obtained by the iteration method.

Table 4

6. Discussion

The value of ratio $(1 - f_0)/(1 - \bar{f})$ is of considerable interest, because it tells us whether or not the population behaves like a panmictic one. I have calculated a large number of numerical examples by formula (4-5) and have found a fact that when the number of colonies is sufficiently large, the value of ratio depends only on the value of mN (number of migrating individuals), and it is almost independent of number of colonies. Furthermore when $mN >> 1$, the value of $(1 - f_0)/(1 - \bar{f})$ is approximately equal to unity. Therefore if a population is of two-dimensional and $mN >> 1$, it behaves very much like a random mating population, even if the population consists of large number of colonies. It is interesting that in a one-dimensional population, the ratio depends on both mN (local property) and the number of colonies (global property), but in a two-dimensional case it depends only on mN (local property). Several examples of one-dimensional and of two-dimensional cases are listed in Table 3.

Table 5.

* In the case of one-dimensional circular population studied Maruyama (1970a,b) in [4, 5], it was shown that if the number of colonies is large, the dominant eigenvalue can be approximated as the smaller of $m\pi^2/n^2$ or $1/nN$, where n = number of colonies. An analogous feature seems to exist in the two-dimensional case studied here. Here we assume that the number of colonies along both axes are the same, i. e., $n_1 = n_2 \equiv n$. If $mN < 1$ the value of $1 - \lambda$ is approximately equal to m/n^2 and if $mN > 2$ it is approximately equal to $1/n^2N$, i. e.,

264

$$1 - \lambda \approx \frac{m}{n^2} \qquad \text{(more precisely } 1 - \lambda < \frac{2m}{n^2}\text{)}$$

$$\text{if } mN < 1$$

$$1 - \lambda \approx \frac{1}{2n^2N} \qquad \text{(more precisely } 1 - \lambda < \frac{1}{n^2N}\text{)}$$

$$\text{if } mN > 1 \quad .$$

This relationship is illustrated in Figure 1.

Let us now assume that a population is divided into partially isolated n colonies of equal size N. Then the inequality $1 - \lambda < 2m/n$ holds for a more general situation. We assume that at each generation colony i exchanges its member with colony j at the rate m_{ij}, and furthermore that the total migration rate $m_T \equiv \sum_{\substack{j \\ j \neq i}} m_{i,j}$ is the same for every colony. The geographical arrangement of colonies and pattern of migration are immaterial. We let $f_{ij}^{(t)}$ be the probability that two randomly chosen homologous genes from colonies i and j are identical by descent, ($f_{ii}^{(t)}$ is the probability for both genes chosen from colony i without replacement), and let $h_{ij}^{(t)} \equiv 1 - f_{ij}^{(t)}$. Then

$$h_{ij}^{(t+1)} = \sum_{k=1}^{n} \sum_{\ell=1}^{n} m_{ki} m_{\ell j} h_{k\ell}^{(t)} - \frac{1}{2N} \sum_{k=1}^{n} m_{ki} m_{kj} h_{kk}^{(t)} \quad .$$

Now let $\underset{\sim}{H}_t \equiv [h_{ij}^{(t)}]_{nxn}$, $\underset{\sim}{M} = [m_{ij}]_{nxn}$ and $\underset{\sim}{G}_t$ be the diagonal

matrix consisting of $h_{11}^{(t)}, h_{22}^{(t)}, \cdots, h_{nn}^{(t)}$. Then the above recurrence relation can be written as

$$\underset{\sim}{H}_{t+1} = \underset{\sim}{M}^T \underset{\sim}{H}_t \underset{\sim}{M} - \frac{1}{2N} \underset{\sim}{M}^T \underset{\sim}{G}_t \underset{\sim}{M}$$

where the superscript T indicates the transposition of matrix.

Let $\| \underset{\sim}{H}_t \| \equiv \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{(t)}/n^2$ and $\| \underset{\sim}{G}_t \| = \sum_{i=1}^{n} h_{ii}^{(t)}/n$.

Note that the matrix $\underset{\sim}{M}$ does not alter the average values, i. e., $\| \underset{\sim}{M}^T \underset{\sim}{H}_t \underset{\sim}{M} \| = \| \underset{\sim}{H}_t \|$ and $\| \underset{\sim}{M}^T \underset{\sim}{G}_t \underset{\sim}{M} \| = \| \underset{\sim}{G}_t \|$. Thus

$$(6\text{-}1) \qquad \| \underset{\sim}{H}_{t+1} \| = \| \underset{\sim}{H}_t \| - \frac{1}{2N} \| \underset{\sim}{G}_t \|$$

This is a generalization of (4-4). Since the form of matrix $\underset{\sim}{H}_t$ does not change asymptotically with time, except it is multiplied by the factor λ, we have the following estimation on $\| \underset{\sim}{G}_t \|$

$$\frac{1}{2N} \| \underset{\sim}{G}_t \| + \left(1 - \frac{1}{2N}\right) \frac{1}{n^2} \sum_{i=1}^{n} \left\{ \sum_{\substack{j,k \\ i \neq j \text{ or } i \neq k}} m_{ij} m_{ik} h_{ii}^{(t)} \right\}$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \left\{ \sum_{\substack{j,k \\ i \neq j \text{ or } i \neq k}} m_{ji} m_{ki} h_{jk} \right\}.$$

Therefore

$$\frac{1}{2N} \| \underset{\sim}{G}_t \| \leq \frac{1}{n^2} \sum_{i=1}^{n} \left\{ \sum_{\substack{j,k \\ i \neq j \text{ or } i \neq k}} m_{ji} m_{ki} h_{jk} \right\}$$

We also note

$$\sum_{\substack{j,k \\ i\neq j \text{ or } i\neq k}} m_{ji}m_{ki}h_{jk} \leq \sum_{\substack{j,k \\ i\neq j \text{ or } i\neq k}} m_{ji}m_{ki}\|H_t\| \quad .$$

Therefore we have

$$\frac{1}{2N}\|G_t\| \leq \frac{\|H_t\|}{n^2}\sum_{i=1}^{n}\left\{\sum_{\substack{j,k \\ i\neq j \text{ or } i\neq k}} m_{ji}m_{ki}\right\} \quad .$$

Ignoring higher order terms of the m_T, we have

$$\frac{1}{2N}\|G_t\| \leq \frac{2m_T\|H_t\|}{n} \quad .$$

Substituting this inequality into (6-1), we have

$$\lambda = \frac{\|H_{t+1}\|}{\|H_t\|} \geq 1 - \frac{2m_T}{n}$$

or

$$1 - \lambda \leq \frac{2m_T}{n}$$

which we wanted to have. The equality is atained if $h_{ii}^{(t)} << h_{ij}^{(t)}$ and $h_{ij}^{(t)} \approx h_{i'j'}^{(t)}$ for every $i \neq j$ and $i' \neq j'$. This seems to be the case in a two-dimensional population with very small amount of migration, i. e., $mN << 1$.

There is a simple but biologically interesting consequency of formula (6-1). Note that $\| \underset{\sim}{G}_t \|$ gives the frequency of heterozygotes at t-th generation. Rearranging the formula slightly and taking sum of $\| \underset{\sim}{G}_t \|$ over all the generations we have

$$\sum_{t=0}^{\infty} \| \underset{\sim}{G}_t \| = 2N \left(\| \underset{\sim}{H}_0 \| - \| \underset{\sim}{H}_\infty \| \right) = 2N \| \underset{\sim}{H}_0 \| ,$$

for $\| \underset{\sim}{H}_\infty \| = 0$. Therefore the expectation of the heterozygotes which appear in the population is invariant under the population structure, and it due to a single mutant gene is equal to twice the population number. This in turn implies that if every mutation occurs at a locus which is homallelic in the entire population, the average probability that a locus in an individual is heterozygous is invariant under the population structure.

SUMMARY

The asymptotic rate of decay of the genetic variability in two-dimensional populations is investigated using the "stepping stone" model in which migration occurs only between geographically adjacent colonies. In a population of n x n colonies each of size N, arranged on the surface of a torus, and with migration rate m, the rate is approximately equal to $m/(2n^2)$ if $mN < 1$, and approximately equal to $1/(2Nn^2)$ if $mN > 1$. When the colonies are arranged on a square plane, the rate is approximately equal to $m/(4n^2)$ if $mN < 2$, and otherwise it is approximately equal to $1/(2Nn^2)$. The criteria, $mN = 1$ (torus-like space) or $mN = 2$ (square plane space), is independent of the habitat size n. An alternative method is employed to obtain a more accurate value of the rate and to obtain the rate for more general models. The associated eigenfunctions are also given.

REFERENCES

Fisher, R. A. 1922. On the dominance ratio. Proc. Roy. Soc. Edinb. 42, 321-341.

Fisher, R. A. 1930. The distribution of gene ratios for rare mutations. Proc. Roy. Soc. Edinb. 50, 205-220.

Kato, T. 1966. Perturbation Theory for Linear Operators. Springer, Berlin.

Maruyama, T. 1970a. On the rate of decrease of heterozygosity in circular stepping models of populations. Theor. Pop. Biol. 1, 101-119.

Maruyama, T. 1970b. Rate of decrease of genetic variability in a subdivided population. Biometrika 57, 299-311.

Maruyama, T. 1971a. The rate of decrease of heterozygosity in a population occupying a circular or a linear habitat. Genetics (~~in press~~). 67, 437-454.

Maruyama, T. 1971b. Closed finite stepping-stone models of population structure. Submitted to J. Appl. Prob.

Robertson, A. 1964. The effect of non-random mating within inbred lines on the rate of inbreeding. Genet. Res. 5, 164-167.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16, 97-159.

Table 1. Numerical value of α from (2-9), and comparisons of the dominant eigenvalue from (2-15) and the exact value obtained by the iteration based on (2-15).

n	N	mN	α	Eigenvalue $(1 - \lambda)$	
				Exact	Approximation by (2-15)
96	2	0.4	1.04	2.07×10^{-5}	1.95×10^{-5}
150	1	0.4	1.00	8.03×10^{-6}	8.00×10^{-6}
100	2	0.5	1.02	2.23×10^{-5}	2.19×10^{-5}
190	1	0.2	1.02	2.79×10^{-6}	2.77×10^{-6}
190	4	0.25	1.02	3.32×10^{-6}	3.24×10^{-6}
190	2	0.3	1.00	3.87×10^{-6}	3.84×10^{-6}

Table 2. Comparison of the dominant eigenvalue calculated from the approximation formula (2-17) and the exact value obtained by the iteration method, for mN > 1.

n	N	mN	Eigenvalue $(1 - \lambda)$	
			Exact	Approximation by (2-17)
190	250	50	5.57×10^{-8}	5.50×10^{-8}
190	15	3	7.75×10^{-7}	7.95×10^{-7}
30	10	3	4.85×10^{-5}	4.78×10^{-5}
96	10	2	4.19×10^{-6}	4.30×10^{-6}
96	20	4	2.35×10^{-6}	2.43×10^{-6}
96	50	10	1.03×10^{-6}	1.04×10^{-6}
10	50	5	9.29×10^{-5}	9.17×10^{-5}
30	50	5	1.01×10^{-5}	1.02×10^{-5}

Table 3. Numerical examples of the eigenvalues obtained by the method of this section, (4-5).

$n_1 \times n_2$	N	mN	$\lim_{u \to 0} \frac{1 - f_0}{1 - \bar{f}}$	Eigenvalue $(1 - \lambda)$	
				from (4-5)	exact value from the iteration
100 x 100	1	0.5	0.447	2.237×10^{-5}	2.231×10^{-5}
150 x 150	1	0.3	0.290	6.451×10^{-6}	6.459×10^{-6}
150 x 150	1	0.4	0.364	8.083×10^{-6}	8.026×10^{-6}
190 x 190	5	0.7	0.456	1.262×10^{-6}	1.259×10^{-6}
190 x 190	10	2.0	0.710	8.931×10^{-7}	8.920×10^{-7}
190 x 190	1	0.3	0.278	3.853×10^{-6}	3.864×10^{-6}
50 x 50	50	5.0	0.889	3.556×10^{-6}	3.554×10^{-6}
50 x 50	10	1.0	0.612	1.223×10^{-5}	1.222×10^{-5}
20 x 20	10	0.1	0.163	2.039×10^{-5}	2.040×10^{-5}
90 x 90	10	1.0	0.578	3.566×10^{-6}	3.570×10^{-6}
20 x 20	10	0.1	0.162	2.019×10^{-5}	2.000×10^{-5}
50 x 50	10	0.1	0.131	2.610×10^{-6}	2.616×10^{-6}
90 x 90	20	4.0	0.852	2.631×10^{-6}	2.632×10^{-6}

Table 4. Comparisons of the eigenfunction obtained from (5-1) and that obtained from the iteration method. Only the diagonal elements of the eigenfunctions are presented, i. e., the values of the h_{ii}. Note that $n_1 = n_2 = n$, * = the values by (5-1) and ** = the exact values by the iteration.

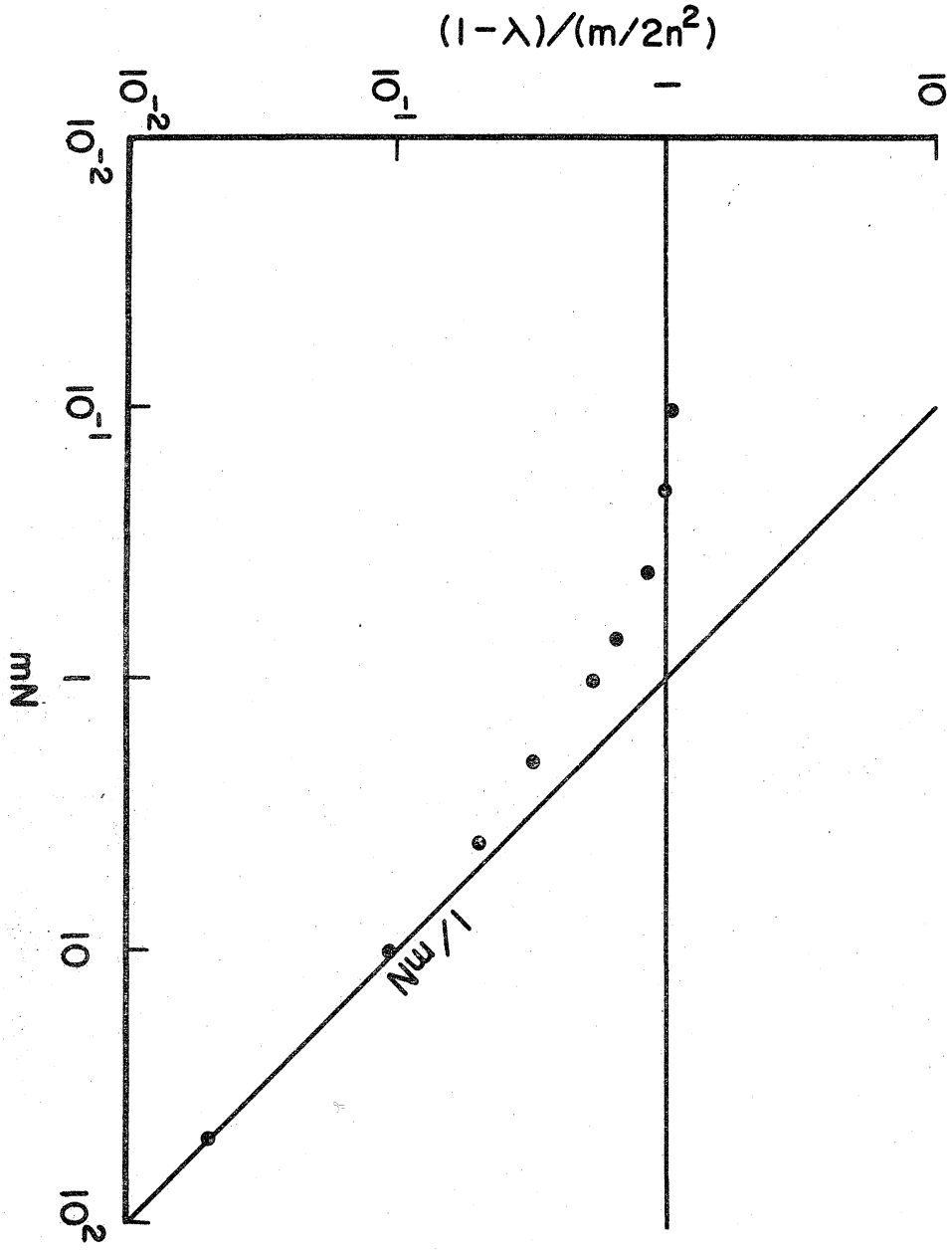
i =	0,	$\frac{n}{20}$,	$\frac{2n}{20}$,	$\frac{3n}{20}$,	$\frac{4n}{20}$,	$\frac{5n}{20}$,	$\frac{6n}{20}$,	$\frac{7n}{20}$,	$\frac{8n}{20}$,	$\frac{9n}{20}$,	$\frac{10n}{20}$,	Method	Parameters
h_{ii} =	.0162	.0666	.0832	.0926	.0986	.1027	.1055	.1074	.1085	.1092	.1094	(*)	n = 20, N = 5
	.0162	.0664	.0831	.0926	.0987	.1028	.1056	.1074	.1086	.1092	.1094	(**)	m = 0.02 ($u = 1.25 \times 10^{-6}$)
h_{ii} =	.0485	.0761	.0861	.0917	.0953	.0977	.0994	.1005	.1012	.1016	.1018	(*)	n = 20, N = 5
	.0485	.0758	.0858	.0915	.0952	.0978	.0995	.1007	.1014	.1018	.1019	(**)	m = 0.1 ($u = 1.25 \times 10^{-6}$)
h_{ii} =	.0074	.0697	.0854	.0940	.0995	.1033	.1059	.1076	.1086	.1092	.1094	(*)	n = 40, N = 5
	.0074	.0693	.0850	.0938	.0995	.1033	.1060	.1077	.1088	.1094	.1096	(**)	m = 0.01 ($u = 1.25 \times 10^{-6}$)

Table 5. Examples of numerical values of $\lim_{u\to 0} \frac{1 - f_0}{1 - \bar{f}}$ obtained from ~~(6)~~ (4-5)

$n_1 \times n_2$	$2N$	mN	$\lim_{u\to 0} \frac{1 - f_0}{1 - \bar{f}}$
100 x 100	20	0.01	0.0127
200 x 200	100	0.01	0.0114
500 x 500	20	0.01	0.0111
200 x 200	100	0.03	0.0329
500 x 500	10	0.05	0.0496
100 x 100	10	0.1	0.1145
300 x 300	10	0.1	0.0955
1000 x 1000	10	0.1	0.0811
100 x 100	20	0.2	0.2054
300 x 300	10	0.2	0.1755
1000 x 1000	10	0.2	0.1588
100 x 100	100	1.0	0.5638
500 x 500	100	1.0	0.5036
200 x 200	40	5.0	0.8603
500 x 500	100	5.0	0.8330
1000 x 1000	40	5.0	0.8242
500 x 500	200	10.0	0.9090
1000 x 1000	200	10.0	0.9008
100 x 100	400	50.0	0.9863
1000 x 1000	200	50.0	0.9809
100 x 1	10	0.5	0.1011
200 x 1	10	0.5	0.0501
500 x 1	10	0.5	0.0208
100 x 1	100	0.5	0.1007
100 x 1	1000	50.0	0.9245
100 x 1	10000	50.0	0.9237

Legend to Figure

Figure 1. Relationship between the rate of steady decay of genetic variation and the migration rate in a two-dimensional stepping stone model. The two lines in the figure represent two approximations (2-15) and (2-17), while the dots represents exact numerical values computed by the matrix iteration method applied to the case of $n^2 = 190 \times 190$ colonies arranged on a torus. N is the effective size of each colony, m is the rate at which each colony exchanges individuals with the 4 sourrounding colonies each generation, and λ is the dominant eigenvalue representing the steady decay. Abscissa: $(1 - \lambda)/(m/2n^2)$. Ordinate: mN.


(1−λ)/(m/2n²)
10⁻²
10⁻¹
1
10
mN
10⁻²
10⁻¹
1
10
10²
Nw/1